

COCOA POLYPEPTIDES AND THEIR USE IN THE PRODUCTION OF COCOA AND CHOCOLATE FLAVOR

5 CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation of International Application PCT/EP02/04258 filed April 17, 2002, the content of which is expressly incorporated herein by reference thereto, which claims priority to EP Application No. 01110251.4 filed April 25, 2001.

10 FIELD OF THE INVENTION

The present invention relates to novel cocoa polypeptides having a molecular weight of about 10 and 14 kDa. The newly identified peptides were first derived from a 69 kDa precursor. In particular, the present invention relates to the production of the polypeptides via recombinant means and the use of the polypeptides or fragments thereof for the production of cocoa/chocolate
15 flavor.

BACKGROUND OF THE INVENTION

The traditional processing of coca beans to generate cocoa flavor requires two steps – a fermentation step, which includes air-drying of the fermented material, and a roasting step.

20 During fermentation the pulp surrounding the beans is degraded by micro-organisms and the sugars contained in the pulp are mainly transformed to acids. In the course of the fermentative process these acids slowly diffuse into the bean eventually causing an acidification of the cellular material. Furthermore, during fermentation peptides of different sizes are generated as well as high levels of hydrophobic free amino acids, which are mainly
25 attributed to the activity of specific proteinases. This specific mixture of peptides and hydrophobic amino acids is thought to be the cocoa-specific flavor precursors.

Research to date has focused on the different proteolytic enzymes involved in these reactions. A number of different types of enzymes, such as an aspartic endoproteinase, a cysteine endoproteinase or a carboxypeptidase have been found to participate in these
30 degradative reactions leading to the formation of the cocoa flavor peptide/amino acid precursor pool.

During the second step of cocoa flavor production – the roasting step – the oligopeptides and amino acids generated during the fermentation stage are subjected to a Maillard reaction in the presence of reducing sugars in the mixture, yielding substances
35 thought to be responsible for the typical cocoa flavor.

There have been attempts to artificially produce cocoa flavor in the past, such as, by subjecting acetone dried powder prepared from unfermented ripe cocoa beans to autolysis at

a pH of 5.2 followed by roasting in the presence of reducing sugars. It was taught that under these conditions preferentially free hydrophobic amino acids and hydrophilic peptides would be generated. The peptide pattern obtained from this process was found to be similar to that of extracts from fermented cocoa beans.

5 Analysis of free amino acids revealed that Leu, Ala, Phe and Val were the predominant amino acids liberated in fermented beans or autolysis (Voigt et al., Food Chem. 49 (1994), 173-180). In contrast to these findings no cocoa-specific flavor could be detected when the above powder was subjected to autolysis at a pH of 3.5. Few free amino acids were found in the by product of the autolysis, but there were a large number of hydrophobic
10 peptides formed.

 Synthetic mixture of free amino acids whose composition resembles that found in fermented beans also have been found to not produce the cocoa flavor desired. These findings indicate that both the peptides and the amino acids are important in producing cocoa flavor (Voigt et al., Food Chem. 49 (1994), 173-180).

15 To date, little attention has been paid to the protein pool from which the peptide/amino acid flavor precursor pool is generated from, since cocoa proteins are often difficult to isolate. One of the major reasons is because that coca seeds contain a high amount of polyphenols and fat. Separating the polyphenols and fat traditionally requires the use of lipophilic organic liquids, such as acetone. The use of these liquids often result in the removal of lipophilic flavor
20 precursors and active substances. Another reason is because of the poor solubilization of proteins purified with acetone, resulting in a poor recovery of the total proteins.

 To date, four major proteins with an apparent molecular weight of 14.5, 21, 31 and 47 kDa, have been identified before fermentation in cocoa bens. These proteins are thought to give rise to the peptide/amino acid pool responsible for producing the cocoa flavor.

25 It is an object of the present invention to further elucidate and identify other protein responsible for producing the cocoa flavor in sufficient detail to eventually provide means for improving the preparation of cocoa flavor and cocoa flavor substitutes.

SUMMARY OF THE INVENTION

30 The present invention is directed to an isolated or synthesized cocoa polypeptide identified by SEQ ID NO:1, SEQ ID NO:2, or a fragment thereof comprising SEQ ID NO:3 or SEQ ID NO:4. These newly identified peptides have a molecular weight of about 10 and 14 kDa, respectively, and were originally derived from the 69 kDa cocoa bean precursor proteins.

35 Typically fragments of the SEQ ID NO:1 and SEQ ID NO:2 are obtained by enzymatic degradation, preferable using one or more of the following enzymes: aspartic endoproteinase, cysteine endoproteinase and carboxypeptidase. In one embodiment the

enzymes used are derived from cocoa plants.

The present invention is also directed to an isolated or synthesized nucleotide sequence that encodes SEQ ID NO:1, SEQ ID NO:2 or fragments thereof. In those embodiments wherein the nucleotide sequence of the invention encodes a fragment of SEQ ID NO:1 and/or SEQ ID NO:2. The peptides encoded by these nucleotides preferably comprises SEQ ID NO:3 and/or SEQ ID NO:4.

The present invention encompasses recombinant cells, vectors, and cells comprising vectors containing one or more copies of the nucleotide sequence described above.

Typically the recombinant cell is a bacterial cell, a yeast cell, an insect cell, a mammalian cell or a plant cell. Preferably the cell is a plant cell and most preferably are part of a plant.

The present invention is further directed to a method of producing cocoa or chocolate flavor comprising isolating, synthesising or producing a polypeptide of the invention. In one embodiment, the method further comprises reacting such a peptide with a reducing sugar.

In yet another embodiment of the invention, the newly identified peptides are used to enhance the cocoa or chocolate flavor of a composition. The method typically comprises supplementing a food composition with one or more of the peptides.

The method can further comprise subjecting the peptide to enzymatic degradation, preferably involving one or more of the following enzymes: aspartic endoproteinase, cysteine endoproteinase or carboxypeptidase, followed by reacting the fragments with reducing sugars.

Still further, the present invention also encompasses a method of producing cocoa beans with increased cocoa flavor proteins. The method typically comprises transforming a cocoa cell with one or more of the nucleotide sequences of the invention followed by generating at least one cocoa plant from the transformed cell. Preferably the transformed cell comprises at least 40 copies of the nucleotide sequence.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows photograph of a two-dimensional SDS-PAGE proteins isolated from unfermented cocoa beans;

Fig. 2 shows the result of a LC_ESI-MS analysis of a GndHCL extract of unfermented cocoa beans.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

During the studies leading to the invention the present inventors have designed novel

methods for an improved isolation of cocoa proteins by using denaturing agents, such as *e.g.* SDS (1%), urea or GndHCl, which resulted in an about 3-fold increase in solubility of proteins as compared to conventional methods. In particular, the use of 6 M GndHCl provided good results. GndHCl showed up to be easily removable by RP-HPLC and no reaction with the proteins occurred. Moreover, it could also be shown that even crude bean powder, after subjection to a treatment with a solubilization buffer including a denaturing agent could be successfully analysed, which made no special care necessary to remove polyphenols.

During the above-described studies directed to provide a better total recovery of cocoa proteins, the inventors ran a crude coca bean powder on a two-dimension SDS-PAGE gel, whereby a cluster of several polypeptides exhibiting a molecular weight of about 9-16 kDa, were detected. The polypeptides contained in the acidic cluster was further isolated by making use of RP-HPLC.

Finally two polypeptides were isolated having a molecular weight of about 10 and about 14 kDa. These polypeptides were N-terminally sequenced and are identified as SEQ ID NO:3 (10 kDa protein; AA residues 1-26 of SEQ ID NO:1) and SEQ ID NO:4 (14 kDa protein; AA residues 1-10 of SEQ ID NO:2), respectively.

Upon comparison with known protein sequences it was shown that these polypeptides were derived from the 69 kDa precursor protein. Upon processing the 69 kDa cocoa beans precursor, it gave rise to the above mentioned 47 and 31 kDa proteins and also to the newly identified 10 and 14 kDa proteins, also representing part of the protein/peptide pool of cocoa beans.

The present invention is directed to an isolated or synthesized cocoa polypeptide identified by SEQ ID NO:1, SEQ ID NO:2. According to another embodiment the present invention, the polypeptides are subjected to enzymatic degradation, preferably with aspartic endoproteinases, cystein endoproteinases and/or carboxypeptidases. In this embodiment it is preferable that the fragments comprise SEQ ID NO:3 or SEQ ID NO:4.

According to yet another embodiment the invention the polypeptides obtainable by the enzymatic degradation are subsequently reacted with reducing sugars.

The present invention also provides for a recombinant nucleotide encoding the polypeptides of the invention, preferably a nucleotide sequence encoding at least one of the new polypeptides or fragments thereof. Such nucleotides may be easily derived from the given polypeptide sequence by translating the amino acid according to the genetic code into corresponding triplets. Such a nucleotide sequence may well be expressed in a suitable cell by means well known in the art, such as *e.g.* in a bacterial cell, *e.g.* in *E. coli*, or in yeast, insect cells, mammalian or plant cells.

To this end, a nucleotide sequence encoding a polypeptide of the present invention is inserted into a suitable vehicle, such as an expression vector, and is incorporated into a cell of

choice. With respect to plant cells, the nucleotides encoding the polypeptides of the present invention may also be incorporated into any of the plant cell's chromosome by using e.g. the phenomenon of homologous recombination. In this respect, at least one copy, preferably more than 40 copies of a nucleotide sequence, encoding any of the present polypeptides may be present on the DNA sequence to be inserted into a plants cell's chromosome.

The present invention further encompasses the generation of plants comprising the recombinant cells. Preferably the transformed plant is a cocoa plant.

Furthermore, the invention provides for the use of the polypeptides for the manufacture of cocoa flavor. To this end it is conceived that the present polypeptides may be added to a fermentation mixture of cocoa beans, in order to provide a higher amount of the polypeptides for degradation. When using cocoa plants that have been modified by recombinant means and contain a high number of copies of nucleotide sequences encoding the polypeptides of the present invention, the plants will inherently contain a higher concentration of the polypeptides and eventually will result in the production of a stronger cocoa flavor after the processing.

EXAMPLES

These and other aspects of the present invention may be more fully understood with reference to the following non-limiting examples, which are merely illustrative of the preferred embodiments of the present invention, and are not to be construed as limiting the invention, the scope of which is defined by the appended claims.

The following examples illustrate the invention in a more detailed manner. It is, however, understood that the present invention is not limited to the examples but is rather embraced by the scope of the appended claims.

Example 1

Separation of proteins

Preparation of crude cocoa bean powder

5 Unfermented cocoa beans were obtained from Ivory Coast and unless stated otherwise all studies were carried out using West African Amelonado cocoa beans. Dried cocoa beans were passed through a bean crusher, followed by a winnover to remove shells. The nibs were kept in a brown bottle at -20 °C. Cocoa nibs were milled for few seconds in an universal mill. The nib powder was passed through 0.8-mm sieve and kept at 4°C.

Two-dimensional SDS-PAGE electrophoretic analysis of unfermented cocoa bean

10 Crude (unfermented) cocoa bean powder (100 mg) was dissolved in 1 ml of solubilization buffer [8 M urea, 3 % (w/v) CHAPS, 2.8 % (v/v) carrier ampholytes (ampholine pH range 4-6.5, 5-8, and 3-10 (2:4:1) and 10 mM DTT (dithiothreitol)]. The clear supernatant was subjected to
15 first dimension of separation on an immobilized pH-gradient (IPG) from 4-7, and second dimension on a 10 % T SDS-PAGE gel. Proteins were visualized by silver staining.

20 The resulting electrophoretic profile of proteins in a typical unfermented cocoa bean on a two-dimensional SDS-PAGE is shown in Fig.1. The 47, 31 and 21 kDa proteins were represented by several subforms and in addition two distinct clusters (acidic and basic) were
25 clearly identified in a molecular weight range of about 9-16 kDa. All of the protein spots could be shown to gradually disappear upon fermentation of beans.

 A Tricine-SDS-PAGE of unfermented cocoa bean genotypes showed up that the clusters in the molecular weight region 9-16 kDa were present in all of the 21 different genotypes representing three cocoa groups, namely Criollo, Forastero and Trinitario.

25 The acidic cluster has been selected for further investigation.

Example 2

Isolation of proteins having a molecular weight of about 9-16 kDa

Preparation of an GndHCl extract of CAP

30 Cocoa nib powder (10g) (supra) was suspended in 200 ml 80 % (v/v) aqueous acetone and stirred for 1 hr at 4 °C. The resulting suspension was centrifuged at 15.000 rpm for 15 min

at 4 °C. The residue was extracted 5-times with 200 ml 80 % (v/v) aqueous acetone and followed by 3-times washing with 100 % acetone. The resulting acetone powder was dried under reduced pressure.

Subsequently a GndHCl extract and a pyridine-ethylated GndHCl extract of CAP from unfermented cocoa beans was prepared. 1 g CAP was suspended with 10 ml GndHCl buffer (100 mM ammonium phosphate, 66.7 mM potassium hydroxide, 3 mM EDTA and 6 M GndHCl) and sonicated for 1 min. The suspension was cooled on ice for 15-30 min and centrifuged at 15000 rpm at 4 °C for 15 min. The clear supernatant was carefully removed. In order to obtain a pyridine-ethylated GndHCl extract, the CAP extract (2 ml) was sparged with argon and mixed with 50 µl of reducing solution (0.8 M DTT in 3 M tripotassium phosphate/3 mM EDTA). The solution was kept at room temperature in dark for 60 min. Pyridine-ethylation at cysteine residues of the reduced CAP was carried out by mixing vigorously 40 µl of 4-vinyl pyridine and further incubation for 30 min at room temperature (Lundell and Schreitmüller, Anal. Biochem. 266 (1999) 31-47). The reaction mixture was dialyzed against 500 ml of the extraction buffer for overnight at room temperature. The dialyzed sample was centrifuged and the clear supernatant passed through 0.22 µm filter disc and kept at 4 °C until analyzation.

LC-ESI-MS analysis of the reduced and pyridylethylated GndHCl-extract

A LC-ESI-MS analysis of the reduced and pyridylethylated extract was performed, as may be seen from FIG. 2. To this end, reduced and pyridylethylated GndHCl extracts of CAP were injected onto reverse phase HPLC columns [Bio-Rad HRLC series 800 system; columns C4 and C8 from Aquapore RP butyl (7 µm, 4.6 x 220 mm), Aquapore RP 300 (7 µm, 4.6 x 220 mm), Perkin-Elmer; Spherisorb 80-5C8 (220 x 4 mm); Marchery Nagel and Vydac protein C4 (4.6 x 220 mm)) pre-equilibrated with solvent A (0.1 % v/v TFA in water) and eluted with a linear gradient of increasing concentration of solvent B (80 % v/v acetonitrile and 0.1 % v/v TFA): 0-15 % B in 5 min, 15-27 % B in 40 min, 27-35 % B in 2 min, isocratic at 35 % B for 3 min, 35-43 % B in 25 min, 43-56 % B in 50 min, 56-70 % B in 5 min, 70-100 % B in 10 min and isocratic at 100 % B for 5 min].

Fractions containing proteins were found to elute at a retention time of about 41, 52, 68, 78 and 87 min, were of about M_r 10,425 (marked as CSP10), 9,010 (marked as CSP9), 20,540 (marked as CSP22) and 12,500 (marked as CSP12), respectively, as can be seen in Table 1 and Fig. 6. In the case of proteins eluting at 41 min and 97 min, no molecular mass could be

identified.

Table 1

Retention time, min	Sample code	Average M_r	Comments
41.4	CSP14	Not detected	
52.5	CSP10	10,425	
67.7	CSP9	9,010	
78.5	CSP22	20,540	Albumin CSP
86.9	CSP12	12,245	
97.3	CSP67	Not detected	Vicilin type CSP
132.2	CSPA _{agg}	Not detected	

Since the average M_w of the protein designated CSP14 could not be assigned with the above method, the peak fraction was dissolved in 500 μ l of 25 % solvent B (0.05 % (v/v) TFA/80 %, v/v ACN). For SDS-PAGE, a 10 μ l aliquot was dried in speedvac and dissolved in 20 μ l SDS-sample buffer and analyzed on gradient 10-20% T ready Tris-Tricine acrylamide gels using the miniprotean 3 system from Bio-Rad. Protein bands were visualized by staining the gels in the staining solution [0.5 % (w/v) Commassie Brilliant Blue R250 in 30 % (v/v) methanol and 10 % (v/v) acetic acid] for 1 hr followed by destaining [30 % (v/v) methanol plus 10 % (v/v) acetic acid] until bands were visible against the clear background (Graffin, Methods Enzymol. 182 (1990) 425-477). Accordingly it could be observed that CSP 14 corresponds to a protein having a molecular weight of about 14 kDa.

Purification/collection by repetitive RP-HPLC

Subsequently, the cocoa proteins, reduced and pyridine-ethylated were isolated/collected by repetitive injections and automatic fraction pooling and collection from the GndHCl extracts of unfermented CAP, as described above. The pooled fractions of each proteins were dried under reduced pressure and dissolved in 400 μ l solvent A and rechromatographed [column Aquapore RP 300 (7 μ m, 4.6 x 220 mm), solvent TFA/ACN system; injection volume 400 μ l; detection at 215 nm; gradients: 1. FIG. 7a (CSP14) and FIG. 7b (CSP9): 0-15 % B in 5 min, isocratic at 15 % B for 5 min, 15-35 % B in 60 min, 35-50 % B in 10 min, 50-100 % B in 5 min and isocratic at 100 % B for 5 min; 2. FIG. 7c (CSP12): 0-35 % B in 5 min, isocratic at 15 % B for 5 min, 35-60 % B in 60 min, 60-75 % B in 10 min, 75-100 % B in 10 min and isocratic at 100 % B for 10 min;]. Fractions 1 ml each were automatically collected and those containing the peak fractions for each of CSP14 and CSP10 were pooled, dried and kept at -20 °C until used.

Example 3

Characterization of purified proteins

The purified cocoa seed proteins CSP10 and CSP14 were subjected to N-terminal amino acid sequencing by automated Edman degradation protein sequencer. The initial and repetitive yield of Edman cycle was between 80 to 90 %. The results obtained are shown in table 2 below.

Table 2

Protein	Initial amount, pmol	Initial yield, pmol	Sequence
CSP10	400	120	RREQE EESEE ETFGE FXQVX APLXP G (SEQ ID NO:3)
CSP14	200	100	GRKQY ERDPR (SEQ ID NO4)

The above listed N-terminal sequence of CSP 10 and 14 has been found to be a part of the 67 kDa vicillin type cocoa storage protein (WO 91/19801, supra). Thus, both of CSP 10 and 14 are so far not identified fragments of the 67 kDa vicilin type cocoa storage protein produced during the normal processing of the protein in cacao beans. By aligning the 47 and 31 kDa proteins, known to be derived from the 67 kDa vicillin protein, to the protein the remaining sequence for the CSP 10 and 14 was derived, which yielded the sequences as identified by SEQ ID NO:1 and SEQ ID NO:2.

A calculation of the molecular weights of the amino acids contained in the polypeptides according to SEQ ID NO:1 and 2 confirmed the approximate molecular weights of the resulting polypeptides were about 10 and 14 kDa.

Consequently the peptides also seem to be excised during the normal processing of the 67 kDa protein and represent a part of the protein/polypeptide pool of cacao beans.